

1644



1600

OCT 22 2002

RECEIVED

TECH CENTER 1600/2800

DATE: 10/07/2002
TIME: 16:56:37

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/500,135C

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

3 <110> APPLICANT: Estell, David A.
 4 Harding, Fiona A.
 6 <120> TITLE OF INVENTION: PROTEINS PRODUCING AN ALTERED IMMUNOGENIC RESPONSE AND
 7 METHODS OF MAKING AND USING THE SAME
 9 <130> FILE REFERENCE: A-68893/DJB/DAV
 11 <140> CURRENT APPLICATION NUMBER: US 09/500,135C
 12 <141> CURRENT FILING DATE: 2000-02-08
 14 <150> PRIOR APPLICATION NUMBER: US 09/060,872
 15 <151> PRIOR FILING DATE: 1998-04-15
 17 <160> NUMBER OF SEQ ID NOS: 236
 19 <170> SOFTWARE: PatentIn Ver. 2.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 1495
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Bacillus amyloliquefaciens
 26 <220> FEATURE:
 27 <221> NAME/KEY: mat_peptide
 28 <222> LOCATION: (417)..(1495)
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (96)..(1244)
 34 <220> FEATURE:
 35 <221> NAME/KEY: misc_feature
 36 <222> LOCATION: (582)..(584)
 37 <223> OTHER INFORMATION: The nnn at positions 582 through 584 which in a
 38 preferred embodiment (aat) is to code for
 39 asparagine, but which may also code for proline.
 41 <220> FEATURE:
 42 <221> NAME/KEY: misc_feature
 43 <222> LOCATION: (585)..(587)
 44 <223> OTHER INFORMATION: The nnn at positions 585 through 587 which in a
 45 preferred embodiment (cct) is to code for proline,
 46 but which may also code for asparagine.
 48 <220> FEATURE:
 49 <221> NAME/KEY: misc_feature
 50 <222> LOCATION: (597)..(599)
 51 <223> OTHER INFORMATION: The nnn at positions 597 to 599 which in a
 52 preferred embodiment (aac) is to code for
 53 asparagine, but which may also code for aspartic acid.
 55 <220> FEATURE:
 56 <221> NAME/KEY: misc_feature
 57 <222> LOCATION: (678)..(680)
 58 <223> OTHER INFORMATION: The nnn at positions 678 through 680 which in a

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/500,135C

DATE: 10/07/2002

TIME: 16:56:37

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

59 preferred embodiment (gca) is to code for
 60 alanine, but which may also code for serine.
 62 <220> FEATURE:
 63 <221> NAME/KEY: misc_feature
 64 <222> LOCATION: (681)..(683)
 65 <223> OTHER INFORMATION: The nnn at positions 681 through 683 which in a
 66 preferred embodiment (tca) is to code for serine,
 67 but which may also code for alanine.
 69 <220> FEATURE:
 70 <221> NAME/KEY: misc_feature
 71 <222> LOCATION: (708)..(710)
 72 <223> OTHER INFORMATION: The nnn at positions 708 through 710 which in a
 73 preferred embodiment (gct) is to code for
 74 alanine, but which may also code for aspartic acid.
 76 <220> FEATURE:
 77 <221> NAME/KEY: misc_feature
 78 <222> LOCATION: (711)..(713)
 79 <223> OTHER INFORMATION: The nnn at positions 711 through 713 which in a
 80 preferred embodiment (gac) is to code for
 81 aspartic acid, but which may also code for alanine.
 83 <220> FEATURE:
 84 <221> NAME/KEY: misc_feature
 85 <222> LOCATION: (888)..(890)
 86 <223> OTHER INFORMATION: The nnn at positions 888 through 890 which in a
 87 preferred embodiment (act) is to code for
 88 threonine, but which may also code for serine.
 90 <220> FEATURE:
 91 <221> NAME/KEY: misc_feature
 92 <222> LOCATION: (891)..(893)
 93 <223> OTHER INFORMATION: The nnn at positions 891 through 893 which in a
 94 preferred embodiment (tcc) is to code for
 95 serine, but which may also code for threonine.
 97 <220> FEATURE:
 98 <221> NAME/KEY: misc_feature
 99 <222> LOCATION: (1167)..(1169)
 100 <223> OTHER INFORMATION: The nnn at positions 1167 through 1169 which in
 101 a preferred embodiment (gaa) is to code for
 102 glutamic acid, but which may also code for glutamine.
 104 <400> SEQUENCE: 1
 105 gqtctactaa aatattattc catactatac aattaataca cagaataatc tgtctattgg 60
 107 ttattctgca aatgaaaaaa aggagaggat aaaga atg aga ggc aaa aaa gta 113
 108 Met Arg Gly Lys Lys Val
 109 -105
 111 tgg atc agt ttg ctg ttt gct tta gcg tta atc ttt acg atg gcg ttc 161
 112 Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu Ile Phe Thr Met Ala Phe
 113 -100 -95 -90
 115 ggc agc aca tcc tct gcc cag gcg gca ggg aaa tca aac ggg gaa aag 209
 116 Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly Lys Ser Asn Gly Glu Lys
 117 -85 -80 -75 -70

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/500,135C

DATE: 10/07/2002
TIME: 16:56:37

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

119	aaa tat att gtc ggg ttt	aaa cag aca atg	agc acg atg	agc gcc gct	257
120	Lys Tyr Ile Val	Gly Phe Lys	Gln Thr Met	Ser Thr Met	Ala Ala
121	-65	-60	-55		
123	aag aag aaa gat	gtc att tct gaa	aaa ggc ggg	aaa gtg caa aag	caa 305
124	Lys Lys Lys	Asp Val Ile Ser	Glu Gly Gly	Lys Val Gln	Lys Gln
125	-50	-45	-40		
127	ttc aaa tat gta	gac gca gct	tca gct aca	tta aac aaa	gct gta 353
128	Phe Lys Tyr Val	Asp Ala Ala Ser	Ala Thr Leu	Asn Glu	Lys Ala Val
129	-35	-30	-25		
131	aaa gaa ttg	aaa aaa gac	ccg agc gtc	gct tac gtt	gaa gaa gat cac 401
132	Lys Glu Leu	Lys Lys Asp	Pro Ser Val	Ala Tyr Val	Glu Glu Asp His
133	-20	-15	-10		
135	gta gca cat	gct tac gcg	cag tcc	gtg cct tac	ggc gta tca caa att 449
136	Val Ala His	Ala Tyr Ala	Gln Ser Val	Pro Tyr Gly Val	Ser Gln Ile
137	-5	-1 1	5	10	
139	aaa gcc cct	gtc ctg cac	tct caa	ggc tac act	gga tca aat gtt aaa 497
140	Lys Ala Pro	Ala Leu His	Ser Gln	Gly Tyr Thr	Gly Ser Asn Val Lys
141	15	20	25		
143	gta gcg gtt	atc gac ago	ggt atc	gat tct tct	cat cct gat tta aag 545
144	Val Ala Val	Ile Asp Ser	Gly Ile Asp	Ser Ser His	Pro Asp Leu Lys
145	30	35	40		
W--> 147	gta gca ggc	gga gcc	agc atg	gtt cct	tct gaa aca nnn nnn ttc caa 593
W--> 148	Val Ala Gly	Gly Ala Ser	Met Val Pro	Ser Glu Thr	Xaa Xaa Phe Gln
149	45	50	55		
W--> 151	gac nnn aac	tct cac gga	act cac	gtt gcc	ggc aca gtt gcg gct ctt 641
W--> 152	Asp Xaa Asn	Ser His Gly	Thr His Val	Ala Gly Thr Val	Ala Ala Leu
153	60	65	70	75	
W--> 155	aat aac tca	atc ggt	gta tta	ggc gtt	gct cca agc nnn nnn ctt tac 689
W--> 156	Asn Asn Ser	Ile Gly Val	Leu Gly Val	Ala Pro Ser	Xaa Xaa Leu Tyr
157	80	85	90		
W--> 159	gct gta	aaa gtt	ctc ggt	nnn nnn ggt	tcc ggc caa tac agc tgg atc 737
W--> 160	Ala Val Lys	Val Leu Gly	Xaa Xaa Gly	Ser Gly Gln	Tyr Ser Trp Ile
161	95	100	105		
163	att aac gga	atc gag	tgg gcg	atc gca	aat atg gac gtt att aac 785
164	Ile Asn Gly	Ile Glu Trp	Ala Ala Asn	Asn Met Asp	Val Ile Asn
165	110	115	120		
167	atg aqc ctc	ggc gga	cct tct	ggt tct	gct tta aaa gcg gca gtt 833
168	Met Ser Leu	Gly Gly Pro	Ser Gly Ser	Ala Ala Leu	Lys Ala Ala Val
169	125	130	135		
171	gat aaa gcc	gtt gca tcc	ggc gtc	gtc gtt	gct gca gcc ggt aac 881
172	Asp Lys Ala	Val Ala Ser	Gly Val Val	Val Ala Ala	Gly Asn
173	140	145	150	155	
W--> 175	gaa ggc	nnn nnn ggc	agc tca	agc aca	gtc tac cct ggt aaa tac 929
W--> 176	Glu Gly Xaa	Xaa Gly Ser	Ser Ser Thr	Val Gly Tyr	Pro Gly Lys Tyr
177	160	165	170		
179	cct tct	gtc att	gca gta	ggc gct	gtt gac agc agc aac caa aga gca 977
180	Pro Ser Val	Ile Ala Val	Gly Ala Val	Asp Ser Ser	Asn Gln Arg Ala
181	175	180	185		
183	tct ttc tca	agc gta	gga cct	gag ctt	gat gtc atg gca cct ggc gta 1025

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/500,135C

DATE: 10/07/2002
TIME: 16:56:37

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

184 Ser Phe Ser Ser Val Gly Pro Glu Leu Asp Val Met Ala Pro Gly Val
185 190 195 200
187 tct atc caa agc acg ctt cct gga aac aaa tac ggg gcg tac aac ggt 1073
188 Ser Ile Gln Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn Gly
189 205 210 215
191 acq tca atg gca tct ccg cac gtt gcc gga gcg gct gct ttg att ctt 1121
192 Thr Ser Met Ala Ser Pro His Val Ala Gly Ala Ala Leu Ile Leu
193 220 225 230 235
W--> 195 tct aag cac ccg aac tgg aca aac act caa gtc cgc agc agt tta nnn 1169
W--> 196 Ser Lys His Pro Asn Trp Thr Asn Thr Gln Val Arg Ser Ser Leu Xaa
197 240 245 250
199 aac acc act aca aaa ctt ggt gat tct ttc tac tat gga aaa ggg ctg 1217
200 Asn Thr Thr Lys Leu Gly Asp Ser Phe Tyr Tyr Gly Lys Gly Leu
201 255 260 265
203 atc aac gta cag gcg gca gct cag taa aacataaaaa accggccttg 1264
204 Ile Asn Val Gln Ala Ala Ala Gln
205 270 275
207 gccccggccgg tttttttatt tttcttcctc cgcatgttca atccgctcca taatcgacgg 1324
209 atqgctccct ctgaaaattt taacgagaaa cggcgggtt acccgctca gtcccgtaac 1384
211 gccccaaatcc tgaaacgtct caatcgccgc ttcccggtt ccggtagct caatgccgta 1444
213 acgggtcgccg gcggtttctt gataccggga gacggcattc gtaatcgat c 1495
216 <210> SEQ ID NO: 2
217 <211> LENGTH: 382
218 <212> TYPE: PRT
219 <213> ORGANISM: Bacillus amyloliquefaciens
220 <220> FEATURE:
221 <221> NAME/KEY: VARIANT
222 <222> LOCATION: (163)...(163)
224 <223> OTHER INFORMATION: Xaa = Asn or Pro
226 <220> FEATURE:
227 <221> NAME/KEY: VARIANT
228 <222> LOCATION: (164)...(164)
229 <223> OTHER INFORMATION: Xaa = Pro or Asn
231 <220> FEATURE:
232 <221> NAME/KEY: VARIANT
233 <222> LOCATION: (168)...(168)
234 <223> OTHER INFORMATION: Xaa = Asn or Asp
236 <220> FEATURE:
237 <221> NAME/KEY: VARIANT
238 <222> LOCATION: (195)...(195)
239 <223> OTHER INFORMATION: Xaa = Ala or Ser
241 <220> FEATURE:
242 <221> NAME/KEY: VARIANT
243 <222> LOCATION: (196)...(196)
244 <223> OTHER INFORMATION: Xaa = Ser or Ala
246 <220> FEATURE:
247 <221> NAME/KEY: VARIANT
248 <222> LOCATION: (205)...(205)
249 <223> OTHER INFORMATION: Xaa = Ala or Asp

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/500,135C

DATE: 10/07/2002
TIME: 16:56:37

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

251 <220> FEATURE:
252 <221> NAME/KEY: VARIANT
253 <222> LOCATION: (206)...(206)
254 <223> OTHER INFORMATION: Xaa = Asp or Ala
256 <220> FEATURE:
257 <221> NAME/KEY: VARIANT
258 <222> LOCATION: (265)...(265)
259 <223> OTHER INFORMATION: Xaa = Thr or Ser
261 <220> FEATURE:
262 <221> NAME/KEY: VARIANT
263 <222> LOCATION: (266)...(266)
264 <223> OTHER INFORMATION: Xaa = Ser or Thr
266 <220> FEATURE:
267 <221> NAME/KEY: VARIANT
268 <222> LOCATION: (358)...(358)
269 <223> OTHER INFORMATION: Xaa = Gln or Glu
271 <400> SEQUENCE: 2
272 Met Arg Gly Lys Lys Val Trp Ile Ser Leu Leu Phe Ala Leu Ala Leu
273 1 5 10 15
274 Ile Phe Thr Met Ala Phe Gly Ser Thr Ser Ser Ala Gln Ala Ala Gly
275 20 25 30
276 Lys Ser Asn Gly Glu Lys Lys Tyr Ile Val Gly Phe Lys Gln Thr Met
277 35 40 45
278 Ser Thr Met Ser Ala Ala Lys Lys Asp Val Ile Ser Glu Lys Gly
279 50 55 60
280 Gly Lys Val Gln Lys Gln Phe Lys Tyr Val Asp Ala Ala Ser Ala Thr
281 65 70 75 80
282 Leu Asn Glu Lys Ala Val Lys Glu Leu Lys Lys Asp Pro Ser Val Ala
283 85 90 95
284 Tyr Val Glu Glu Asp His Val Ala His Ala Tyr Ala Gln Ser Val Pro
285 100 105 110
286 Tyr Gly Val Ser Gln Ile Lys Ala Pro Ala Leu His Ser Gln Gly Tyr
287 115 120 125
288 Thr Gly Ser Asn Val Lys Val Ala Val Ile Asp Ser Gly Ile Asp Ser
289 130 135 140
290 Ser His Pro Asp Leu Lys Val Ala Gly Gly Ala Ser Met Val Pro Ser
291 145 150 155 160
W--> 292 Glu Thr Xaa Xaa Phe Gln Asp Xaa Asn Ser His Gly Thr His Val Ala
293 165 170 175
294 Gly Thr Val Ala Ala Leu Asn Asn Ser Ile Gly Val Leu Gly Val Ala
295 180 185 190
W--> 296 Pro Ser Xaa Xaa Leu Tyr Ala Val Lys Val Leu Gly Xaa Xaa Gly Ser
297 195 200 205
298 Gly Gln Tyr Ser Trp Ile Ile Asn Gly Ile Glu Trp Ala Ile Ala Asn
299 210 215 220
300 Asn Met Asp Val Ile Asn Met Ser Leu Gly Gly Pro Ser Gly Ser Ala
301 225 230 235 240
302 Ala Leu Lys Ala Ala Val Asp Lys Ala Val Ala Ser Gly Val Val Val
303 245 250 255

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 10/07/2002
PATENT APPLICATION: US/09/500,135C TIME: 16:56:38

Input Set : A:\GC527-C1-revseqlist.txt
Output Set: N:\CRF4\10072002\I500135C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 582,583,584,585,586,587,597,598,599,678,679,680,681,682,683

Seq#:1; N Pos. 708,709,710,711,712,713,888,889,890,891,892,893,1167,1168

Seq#:1; N Pos. 1169

Seq#:1; Xaa Pos. 56,57,61,88,89,98,99,158,159,251

Seq#:2; Xaa Pos. 163,164,168,195,196,205,206,265,266,358